PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/08/785,532*

SECT 1816

DATE: 07/10/97
TIME: 18:42:55

INPUT SET: S18901.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

```
 1                              SEQUENCE LISTING
 2
 3   (1)     General Information:
 4
 5       (i) APPLICANT: Gray, Joe W.
 6                      Collins, Colin
 7                      Hwang, Soo-In
 8                      Godfrey, Tony
 9                      Kowbel, David
10                      Rommens, Johanna
11
12      (ii) TITLE OF INVENTION: Genes From the 20q13 Amplicon and Their
13           Uses
14
15     (iii) NUMBER OF SEQUENCES: 55
16
17      (iv) CORRESPONDENCE ADDRESS:
18           (A) ADDRESSEE: Townsend and Townsend and Crew LLP
19           (B) STREET: Two Embarcadero Center, 8th Floor
20           (C) CITY: San Francisco
21           (D) STATE: California
22           (E) COUNTRY: USA
23           (F) ZIP: 94111-3834
24
25       (v) COMPUTER READABLE FORM:
26           (A) MEDIUM TYPE: Floppy disk
27           (B) COMPUTER: IBM PC compatible
28           (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29           (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
30
31      (vi) CURRENT APPLICATION DATA:
32           (A) APPLICATION NUMBER: US 08/785,532
33           (B) FILING DATE: 17-JAN-1997
34           (C) CLASSIFICATION:
35
36     (vii) PRIOR APPLICATION DATA:
37           (A) APPLICATION NUMBER: US 08/731,499
38           (B) FILING DATE: 16-OCT-1996
39
40     (vii) PRIOR APPLICATION DATA:
41           (A) APPLICATION NUMBER: US 08/680,395
42           (B) FILING DATE: 15-JUL-1996
43
44    (viii) ATTORNEY/AGENT INFORMATION:
45           (A) NAME: Quine, Jonathan A.
46           (B) REGISTRATION NUMBER: P-41,261
```

ENTERED

```
47              (C) REFERENCE/DOCKET NUMBER: 023070-068920US
48
49        (ix) TELECOMMUNICATION INFORMATION:
50              (A) TELEPHONE: (415) 576-0200
51              (B) TELEFAX: (415) 576-0300
52
53
54  (2) INFORMATION FOR SEQ ID NO:1:
55
56        (i) SEQUENCE CHARACTERISTICS:
57              (A) LENGTH: 3000 base pairs
58              (B) TYPE: nucleic acid
59              (C) STRANDEDNESS: single
60              (D) TOPOLOGY: linear
61
62       (ii) MOLECULE TYPE: cDNA
63
64
65       (ix) FEATURE:
66              (A) NAME/KEY: -
67              (B) LOCATION: 1..3000
68              (D) OTHER INFORMATION: /note= "cDNA clone 3bf4 of 3kb
69                                   transcript of tyrosine kinase gene A6"
70
71       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
72
73  CCGCCGGCCG GGGCGCCTGG CTGCACTCAG CGCCGGAGCC GGGAGCTAGC GGCCGCCGCC      60
74
75  ATGTCCCACC AGACCGGCAT CCAAGCAAGT GAAGATGTTA AAGAGATCTT TGCCAGAGCC     120
76
77  AGAAATGGAA AGTACAGACT TCTGAAAATA TCTATTGAAA ATGAGCAACT TGTGATTGGA     180
78
79  TCATATAGTC AGCCTTCAGA TTCCTGGGAT AAGGATTATG ATTCCTTTGT TTTACCCCTG     240
80
81  TTGGAGGACA AACAACCATG CTATATATTA TTCAGGTTAG ATTCTCAGAA TGCCCAGGGA     300
82
83  TATGAATGGA TATTCATTGC ATGGTCTCCA GATCATTCTC ATGTTCGTCA AAAAATGTTG     360
84
85  TATGCAGCAA CAAGAGCAAC TCTGAAGAAG GAATTTGGAG GTGGCCACAT TAAAGATGAA     420
86
87  GTATTTGGAA CAGTAAAGGA AGATGTATCA TTACATGGAT ATAAAAAATA CTTGCTGTCA     480
88
89  CAATCTTCCC CTGCCCCACT GACTGCAGCT GAGGAAGAAC TACGACAGAT TAAAATCAAT     540
90
91  GAGGTACAGA CTGACGTGGG TGTGGACACT AAGCATCAAA CACTACAAGG AGTAGCATTT     600
92
93  CCCATTTCTC GAGAAGCCTT TCAGGCTTTG GAAAAATTGA ATAATA●ACA GCTCAACTAT     660
94
95  GTGCAGTTGG AAATAGATAT AAAAAAATGAA ATTATAATTT T●GCCAACAC AACAAATACA     720
96
97  GAACTGAAAG ATTTGCCAAA GAGGATTCCC AAGGATTCAG C●CGTTACCA TTTCTTTCTG     780
98
99  TATAAACATT CCCATGAAGG AGACTATTTA GAGTCCATAG TTTTTATTTA TTCAATGCCT     840
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/785,532*

INPUT SET: S18901.raw

```
100
101   GGATACACAT GCAGTATAAG AGAGCGGATG CTGTATTCTA GCTGCAAGAG CCGTCTGCTA      900
102
103   GAAATTGTAG AAAGACAACT ACAAATGGAT GTAATTAGAA AGATCGAGAT AGACAATGGG      960
104
105   GATGAGTTGA CTGCAGACTT CCTTTATGAA GAAGTACATC CCAAGCAGCA TGCACACAAG     1020
106
107   CAAAGTTTTG CAAAACCAAA AGGTCCTGCA GGAAAAAGAG GAATTCGAAG ACTAATTAGG     1080
108
109   GGCCCAGCGG AAACTGAAGC TACTACTGAT TAAAGTCATC ACATTAAACA TTGTAATACT     1140
110
111   AGTTTTTTAA AAGTCCAGCT TTTAGTACAG GAGAACTGAA ATCATTCCAT GTTGATATAA     1200
112
113   AGTAGGGAAA AAAATTGTAC TTTTTGGAAA ATAGCACTTT TCACTTCTGT GTGTTTTTAA     1260
114
115   AATTAATGTT ATAGAAGACT CATGATTTCT ATTTTTGAGT TAAAGCTAGA AAAGGGTTCA     1320
116
117   ACATAATGTT TAATTTTGTC ACACTGTTTT CATAGCGTTG ATTCCACACT TCAAATACTT     1380
118
119   CTTAAAATTT TATACAGTTG GGCCAGTTCT AGAAAGTCTG ATGTCTCAAA GGGTAAACTT     1440
120
121   ACTACTTTCT TGTGGGACAG AAAGACCTTA AAATATTCAT ATTACTTAAT GAATATGTTA     1500
122
123   AGGACCAGGC TAGAGTATTT TCTAAGCTGG AAACTTAGTG TGCCTTGGAA AAGCCGCAAG     1560
124
125   TTGCTTACTC CGAGTAGCTG TGCTAGCTCT GTCAGACTGT AGGATCATGT CTGCAACTTT     1620
126
127   TAGAAATAGT GCTTTATATT GCAGCAGTCT TTTATATTTG ACTTTTTTTT AATAGCATTA     1680
128
129   AAATTGCAGA TCAGCTCACT CTGAAACTTT AAGGGTACCA GATATTTTCT ATACTGCAGG     1740
130
131   ATTTCTGATG ACATTGAAAG ACTTTAAACA GCCTTAGTAA ATTATCTTTC TAATGCTCTG     1800
132
133   TGAGGCCAAA CATTTATGTT CAGATTGAAA TTTAAATTAA TATCATTCAA AAGGAAACAA     1860
134
135   AAAATGTTGA GTTTTAAAAA TCAGGATTGA CTTTTTTCTC CAAAACCATA CATTTATGGG     1920
136
137   CAAATTGTGT TCTTTATCAC TTCCGAGCAA ATACTCAGAT TTAAAATTAC TTTAAAGTCC     1980
138
139   TGGTACTTAA CAGGCTAACG TAGATAAACA CCTTAATAAT CTCAGTTAAT ACTGTATTTC     2040
140
141   AAAACACATT TAACTGTTTT CTAATGCTTT GCATTATCAG TTACAACCTA GAGAGATTTT     2100
142
143   GAGCCTCATA TTTCTTTGAT ACTTGAAATA GAGGGAGCTA GAACACTTAA TGTTTAATCT     2160
144
145   GTTAAACCTG CTGCAAGAGC CATAACTTTG AGGCATTTTC TAAATGAACT GTGGGGATCC     2220
146
147   AGGATTTGTA ATTTCTTGAT CTAAACTTTA TGCTGCATAA ATCACTTATC GGAAATGCAC     2280
148
149   ATTTCATAGT GTGAAGCACT CATTTCTAAA CCTTATTATC TAAGGTAATA TATGCACCTT     2340
150
151   TCAGAAATTT GTGTTCGAGT AAGTAAAGCA TATTAGAATA ATTGTGGGTT GACAGATTTT     2400
152
```

INPUT SET: S18901.raw

```
153    TAAAATAGAA TTTAGAGTAT TTGGGGTTTT GTTTGTTTAC AAATAATCAG ACTATAATAT    2460
154
155    TTAAACATGC AAAATAACTG ACAATAATGT TGCACTTGTT TACTAAAGAT ATAAGTTGTT    2520
156
157    CCATGGGTGT ACACGTAGAC AGACACACAT ACACCCAAAT TATTGCATTA AGAATCCTGG    2580
158
159    AGCAGACCAT AGCTGAAGCT GTTATTTTCA GTCAGGAAGA CTACCTGTCA TGAAGGTATA    2640
160
161    AAATAATTTA GAAGTGAATG TTTTTCTGTA CCATCTATGT GCAATTATAC TCTAAATTCC    2700
162
163    ACTACACTAC ATTAAAGTAA ATGGACATTC CAGAATATAG ATGTGATTAT AGTCTTAAAC    2760
164
165    TAATTATTAT TAAACCAATG ATTGCTGAAA ATCAGTGATG CATTTGTTAT AGAGTATAAC    2820
166
167    TCATCGTTTA CAGTATGTTT TAGTTGGCAG TATCATACCT AGATGGTGAA TAACATATTC    2880
168
169    CCAGTAAATT TATATAGCAG TGAAGAATTA CATGCCTTCT GGTGGACATT TTATAAGTGC    2940
170
171    ATTTTATATC ACAATAAAAA TTTTTTCTCT TTAAAAAAAA AAAACAAGAA AAAAAAAAA     3000
172
173
174    (2) INFORMATION FOR SEQ ID NO:2:
175
176         (i) SEQUENCE CHARACTERISTICS:
177             (A) LENGTH: 723 base pairs
178             (B) TYPE: nucleic acid
179             (C) STRANDEDNESS: single
180             (D) TOPOLOGY: linear
181
182        (ii) MOLECULE TYPE: cDNA
183
184
185        (ix) FEATURE:
186             (A) NAME/KEY: -
187             (B) LOCATION: 1..723
188             (D) OTHER INFORMATION: /note= "cDNA clone 1b11 of 3.5kb
189                                       transcript"
190
191
192        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
193
194    TGGAAGCTGT CATGGTTACC GTCTCTAACG TTGGACTCTT AAGAAAATGA TTATTCCTGG      60
195
196    TTTCTAGACA GGCCAAATGT AATTCACCTA CGTGGCAGAT TAAAGAGGTG GGCTTACTAG     120
197
198    ATTTGATTGG GTATTGAGCA TGCTCTGAAT GACAGTCCCC AAAAAGGACC TCTTATCCGT     180
199
200    TCTTCCCCTT GGGGAAGGGC TTTTGCCACT TCCATGTCAA TGTGGCAGTT GAGCTTGGAA     240
201
202    ATTGGTGCGT TGTACAACAT AAGCATTACT TCTCCAAGAT GTGCCTGTGT AGAAATGGTC     300
203
204    ATAGATTCAA AACTGTAGCT ACTATGTGGA CAGGGGGGCA GCAAGGACCC CACTTTGTAA     360
205
```

INPUT SET: S18901.raw

```
206   AACATGTTTT GGGGGAATGT TTTGTTTTTC ATTTTCTTAT TACCTGGCAA AATAATCCAG      420
207
208   GTGGTGTGTG AGTCACCAGT AGAGATTATA AAGTCCAAGG AAGTAGAATC AGCCTTACAA      480
209
210   ACAGTGGACC TCAACGAAGG AGATGCTGCA CCTGAACCCA CWGAAGCGAA ACTCAAAAGA      540
211
212   GAAGAAAGCA AACCAAGAAC CTCTCTGATG RCGTTTCTCA GACAAATGGT AAGCCCCTTA      600
213
214   CTTCCAGTAT AGGAAACCTA AGATACCTAG AGCGGCTTTT GGGAACAATG GGCTCATGCC      660
215
216   ACAGGTAGTA GGAGACATAA TTGTAGCTGG TGTGTATGGA ATGTGAATGG AATATGGATT      720
217
218   GCG                                                                   723
219
220
221   (2) INFORMATION FOR SEQ ID NO:3:
222
223       (i) SEQUENCE CHARACTERISTICS:
224           (A) LENGTH: 1507 base pairs
225           (B) TYPE: nucleic acid
226           (C) STRANDEDNESS: single
227           (D) TOPOLOGY: linear
228
229   (ii) MOLECULE TYPE: cDNA
230
231
232   (ix) FEATURE:
233           (A) NAME/KEY: -
234           (B) LOCATION: 1..1507
235           (D) OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
236                                 transcript with homology to C2H2 zinc
237                                 finger genes"
238
239
240   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
241
242   GCAGGTTGCT GGGATTGACT TCTTGCTCAA TTGAAACACT CATTCAATGG AGACAAAGAG       60
243
244   CACTAATGCT TTGTGCTGAT TCATATTTGA ATCGAGGCAT TGGGAACCCT GTATGCCTTG      120
245
246   TTTGTGGAAA GAACCAGTGA CACCATCACT GAGCTTCCTA AAAGTTCGAA GAAGTTAGAG      180
247
248   GACTATACAC TTTCTTTTGA ACTTTTATAA TAAATATTTG CTCTGGTTTT GGAACCCAGG      240
249
250   ACTGTTAGAG GGTGAGTGAC AGGTCTTACA GTGGCCTTAA TCCAACTCCA GAAATTGCCC      300
251
252   AACGGAACTT TGAGATTATA TGCAATCGAA AGTGACAGGA AACATGCCAA GTCAATCCCT      360
253
254   CTTAATGTAC ATGGATGGCC AAGAGTGATT GGCAGCTCTC TTGCCAGTCC GATGGAGATG      420
255
256   GAGATGCCTT GTCAATGAAA GGGCCCNCTG TTGTCAATTC CGAGCTACAC AAAGAAAAAA      480
257
258   ATGTCAATCC GAATCGAGGG GAATATGCCC TTGGATTGCA TGTTCTGCAG CCAGACCTTC      540
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/785,532*

INPUT SET: S18901.raw

Line   Error   Original Text